

Classification of Land-Cover Types for the Fort Benning Ecoregion Using Enhanced Thematic Mapper Data

by Scott G. Bourne, Mark R. Graves

PURPOSE: Information regarding regional land cover is a fundamental requirement to support the long-term baseline ecosystem monitoring plan under the Strategic Environmental Research and Development Program (SERDP), Ecosystem Management Project (SEMP), Ecosystem Characterization and Monitoring Initiative (ECMI). The land cover characterization phase of this plan provides the foundation needed to derive vegetation density indices and land cover patterns. These characteristics are the primary visible expressions of the underlying ecosystem structure, function, and process at all spatial scales (Kress 2000). To meet the requirement for land cover information, Landsat 7 Enhanced Thematic Mapper (ETM) data were used to classify land cover types for the Fort Benning ecoregion. This technical note (TN) describes the procedures used to extract land cover information from the satellite imagery.

BACKGROUND: At a regional scale, land cover significantly affects biophysical factors such as surface albedo and sensible heat flux, and plays an important role in material cycling. Developing an accurate land cover classification is vital, since other landscape characteristics are directly linked to it.

Satellite imagery has been used since the 1970's as an accurate and cost-effective tool for deriving regional vegetation and land cover information. Digital processing techniques involving the statistical analysis of image data representing various portions of the electromagnetic spectrum allow definition of areas that reflect solar radiation in a like manner. These areas may then be related to land cover or vegetation types through the use of ground-truth data collected in the field.

Landsat 7 ETM data were selected for this study due to the spectral and spatial characteristics of the sensor, which have been documented as appropriate data for mapping broad vegetative types such as deciduous and evergreen forests (Schriever and Congalton 1993). In addition, Landsat 7 data are relatively inexpensive and the scenes required to cover the study area could be acquired quickly. The Landsat 7 satellite ETM sensor provides six spectral bands of imagery, each with a spatial resolution of 25 m. The ETM sensor also provides one panchromatic (black and white) band with 15-m spatial resolution and one thermal band with 60-m spatial resolution.

ETM data collected on July 24, 1999 (scene LE7019037009920550 path 019 / row 37 and scene LE7019038009920550 path 019 / row 038) were used in the study. This date of imagery was selected to allow optimum discrimination of vegetation types.

STUDY AREA: Fort Benning, GA, is located in west central Georgia, south of the city of Columbus, Georgia, and east of Phenix City, Alabama. The study area is comprised of approximately 180,687 ha, which includes the military base and a portion of the surrounding U.S.

Geological Survey (USGS) Hydrologic Unit Code (HUC) 03130003. The base covers 74,370 ha. Figure 1 illustrates location of the military installation and HUC. The land cover of the study area is dominated by a variety of evergreen and deciduous forest types distributed over moderate rolling topography.

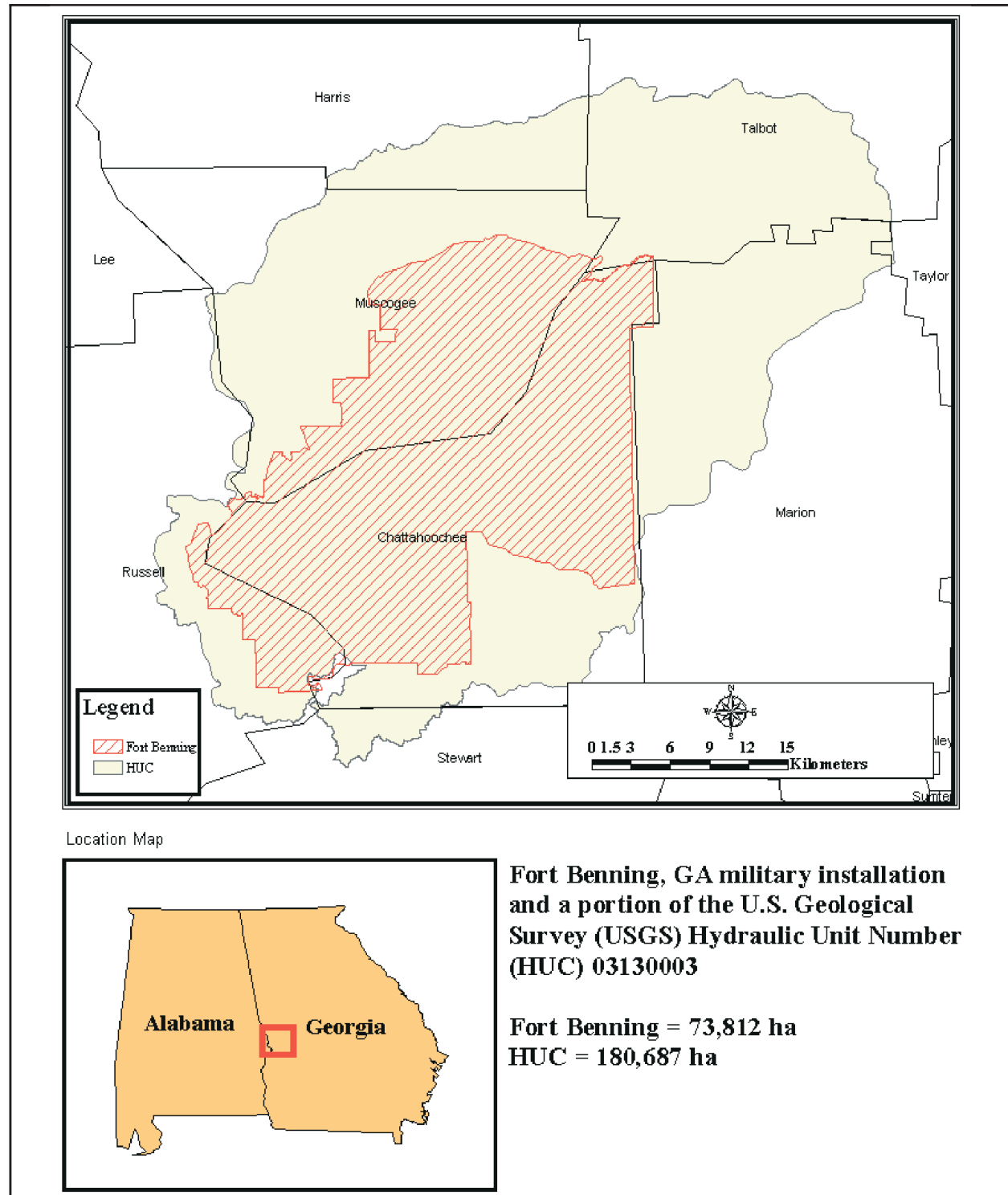


Figure 1. Location of study area

METHOD: ERDAS Imagine (Version 8.4) software was used to perform all image processing functions required to complete the land cover classification. A hybrid supervised/unsupervised classification methodology was used. In a supervised classification, the image analyst is responsible for defining “training areas,” which represent all of the cover types he or she wishes to extract from an image. However, it is often difficult to account for all the cover types in an image as well as variability within cover types. Unsupervised classification differs from supervised classification in that the computer (rather than the user) develops the signatures that will be used to classify the scene. The classification process results in a number of spectral classes, which the analyst must then assign (a posteriori) to information classes of interest. This requires knowledge of the terrain present in the scene as well as its spectral characteristics.

Fleming, Berkebile, and Hoffer (1975) outlined a hybrid approach for image classification that makes use of the benefits of both supervised and unsupervised approaches. This approach requires the following four steps:

- 1) Use clustering algorithms to determine the spectral classes into which the image can be divided.
- 2) Use ground truth data to assign information classes to the statistical clusters.
- 3) Use statistical distance measures to evaluate the initial clusters. Delete or merge clusters as necessary.
- 4) Using a maximum-likelihood algorithm, classify the entire image into the set of spectral classes.

To generate a better set of class statistics for the bare ground land cover type, cantonment areas and paved roads were removed from the Landsat digital image before the unsupervised clustering algorithm was used. These land cover types have been mapped and are available in the Fort Benning Geographical Information System (GIS) database. This would produce a better training class for bare ground by eliminating any chance of confusing bare ground with paved roads or features in the cantonment area that have similar reflectance properties.

Definition of Initial Clusters. In unsupervised classifications, statistical clustering algorithms are used to analyze the digital values in each band of imagery and to determine the number of statistically distinct features (clusters) in the image. In this study, an unsupervised iterative self-organizing data analysis (ISODATA) clustering algorithm was used. ISODATA is a widely used clustering algorithm that makes a large number of passes through an image using a minimum spectral distance formula to form clusters. It begins with either arbitrary cluster means or means of an existing signature set, and each time the clustering repeats, the means of these clusters are shifted. The new cluster means are used for the next iteration. This iterative process continues until statistically distinct features emerge.

To perform ISODATA clustering, it was necessary to specify the Landsat bands to be used for the classification. For this study the three visible bands (TM1, TM2, and TM3), one near infrared band (TM4), and two middle infrared bands (TM5 and TM7) were used. Next, the maximum numbers of clusters are determined. Fifty initial clusters were requested. Such a large number of clusters were requested to provide a wide variation of land cover types that could be easily identified.

Requesting too few clusters could have caused the ISODATA clustering method to combine different land cover types. Since each cluster is the basis for a class, this number becomes the maximum number of classes to be formed.

The ISODATA process began by determining the 50 arbitrary cluster means that were requested. On the first iteration of the ISODATA algorithm, the means of the 50 clusters were arbitrarily determined. After each iteration, a new mean for each cluster was calculated, based on the actual spectral locations of the pixels in the cluster, instead of the initial arbitrary calculation. These new means were then used for defining clusters in the next iteration. The process continued until there was little change between iterations. The convergence threshold was set to 95 percent, which is the maximum percentage of pixels whose class values are allowed to be unchanged between iterations. After each iteration, the normalized percentage of pixels whose assignments are unchanged since the last iteration is displayed, and when the percentage of unchanged pixels reaches 95 percent, the classification is completed (Smith, Pyden, and Cole 1995).

Evaluation of Clusters: Statistical separability tools were used to determine if clusters from ISODATA were to be used as a class in the final classification, combined with another class generating a new land cover type, or discarded. Signature separability is a statistical measure of distance between two signatures. Separability can be calculated for any combination of bands that will be used in the classification generated by the ISODATA clustering algorithm. Bands 4 and 5 were used to calculate the separability of the class signatures. Imagine 8.4 software uses these bands as the default bands.

Transformed divergence (TD) is a measure of separability that was used to determine whether the signatures were separable. TD values have an upper (2000) and a lower bound (0). If the calculated divergence is equal to the upper bound, then the signatures can be said to be totally separable in the bands being analyzed. A calculated divergence of zero means that the signatures are inseparable. For this study, any class combination with a TD above 1500 was considered separable. Any class combination with a TD below 1500 was combined or one of the classes was disregarded. The next step was to select the class combination with the lowest TD. Each class was analyzed individually to identify the class type (i.e., hardwood forest, bare ground, etc.) through the use of ground truth data. After the class types were determined, the TD separability values were used to decide which class types should be deleted or combined. For example, signature 4 and signature 8 had a TD separability of 987. Signature 8 also had separability values lower than 1500 with three other class signatures, while signature 4 was separable from all other class signatures. Signature 4 was considered a better class and signature 8 was therefore deleted. An analysis of the number of pixels in each class and the distribution of pixels in the classified images is accomplished before generating the final classification.

After analyzing class combinations with a TD of less than 1500, the final number of separable classes was 23. Each of the final 23 classes was found to belong to one of the following 10 classes (there was more than one statistically separable type of hardwood, for example):

Table 1 Land Cover Types	
Water	Scrub/Shrub
Evergreen Planted	Herbaceous
Evergreen	Bare Ground
Hardwood	Paved Roads
Cantonment	Not Mapped

Maximum-Likelihood Classification: The next stage of the classification process involved using a maximum-likelihood algorithm. The 23 final class signatures developed using ISODATA were input to the maximum-likelihood algorithm. The maximum-likelihood classification makes use of the statistical parameters developed through the ISODATA process and, in addition, uses estimates of probability distributions to determine the relative likelihood that a pixel belongs to a certain class.

Even though these classes had been identified once using ISODATA, each class was checked with the DOQ's and Forest Stand coverage to make sure that classes were mapped correctly.

Refinement of Final Product: Once the maximum-likelihood classification was complete, there were many isolated, very small classes present. A filtering approach was used to eliminate the small areas that occurred through the classified image. These small areas are one to nine pixels in size. A 3×3 low-pass filter was used to eliminate these areas. The first application of the filter removes any small areas in the classified image. The second application eliminates any small areas generated by using the filter the first time.

Using the capabilities of the Imagine software, layers in the Fort Benning GIS database representing cantonment areas, paved roads, and tank trails were merged onto the final classification.

Results: During the classification process, 50 training classes were initially generated from the Landsat image. Classes with spectral similarities were then aggregated to derive a land cover type map with 23 classes. The 23 classes were then further combined to form a final land cover map with 9 unique classes. Final classification results for the entire HUC and the military base are summarized in Tables 2 and 3, respectively. The land cover classification is displayed in Figure 2.

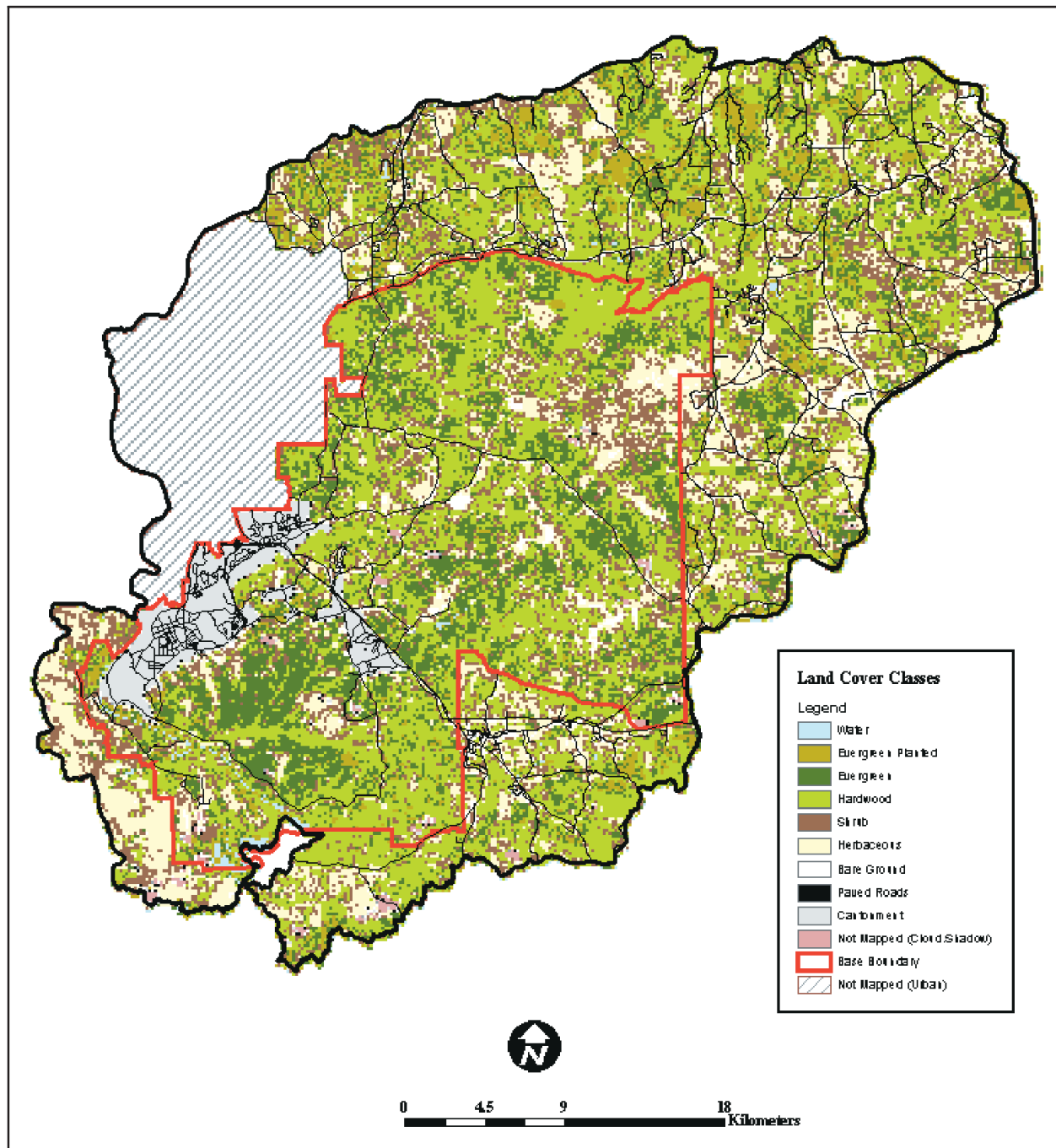


Figure 2. Land cover classification for Fort Benning Ecoregion

Table 2 Unsupervised Classification Results of the HUC			
Name	Area in Hectares	Area in Acres	% of Total
Water	1,949	4,817	1.1%
Evergreen Planted	16,895	41,748	9.4%
Evergreen	25,342	62,622	14.0%
Hardwood	59,249	146,519	32.8%
Scrub/Shrub	21,898	54,112	12.1%
Herb	26,024	64,308	14.4%
Bare Ground	2,274	5,620	1.1%
Paved Roads	3,476	8,590	0.2%
Cantonment	5,426	13,407	1.9%
Not Mapped	18,107	44,743	3.0%
Total	180,687	446,487	100.0%

Table 3 Unsupervised Classification Results of the Fort Benning Military Installation			
Name	Area in Hectares	Area in Acres	% of Total
Water	714	1,766	1.0%
Evergreen Planted	3,801	9,393	5.1%
Evergreen	19,721	48,733	26.5%
Hardwood	26,056	64,388	35.0%
Scrub/Shrub	9,227	22,801	12.4%
Herb	6,206	15,334	8.3%
Bare Ground	1,332	3,292	1.7%
Paved Roads	1,300	3,211	1.7%
Cantonment	5,426	13,407	7.3%
Not Mapped	585	1,455	0.8%
Total	74,370	183,771	100.0%

As with all land cover maps produced using this process, the result of the classification represents only a generalization of the real landscape types. Therefore, it is necessary to check the accuracy of the land cover classification with ground truth data, if available (Ahmad, Jupp, and Nunez 1992). Ground truth data and the classification results must be compared and statistically analyzed. Contingency tables or error matrices are common statistical methods used to compare the results obtained from the classification. An error matrix analysis provides a natural framework for the convenient display of results that can also be used for analysis. This is an effective tool that presents the overall accuracy of the classification as well as the producer and user accuracy of each category (Congalton, Oderwald, and Mead 1983).

The classification accuracy assessment was based on 187 Land Condition — Trend Analysis (LCTA) plot data obtained from Fort Benning. Vegetated plots are identified by specific vegetation types, which follow the National Plant List of Scientific Names codes. From these codes, a general vegetation type was assigned to each LCTA plot. The general vegetation types assigned to each LCTA plot match the classification scheme in Table 1.

The level of agreement or disagreement between the Landsat classification and the reference data is shown in Table 4. The overall accuracy of the classification is 69.5 percent. The overall accuracy is calculated by summing the main diagonal elements of the error matrix and dividing by the total number of samples. The major forest stand classes (hardwood and evergreen) displayed the highest level of agreement with the reference data, with user accuracies of 85 and 83 percent, respectively. Data were not available to check the accuracy of the water, paved roads, and cantonment land cover types.

Table 4 Error Matrix												
Reference Data	Classification	Water	Evergreen/Planted	Evergreen	Hardwood	Scrub/Shrub	Herbaceous	Bare Ground	Paved Roads	Cantonment	Row Total	Producer's Accuracy %
	Water	0	0	0	0	0	0	0	0	0	0	0
	Evergreen/Planted	0	14	1	2	1	2	0	0	0	20	70
	Evergreen	0	4	35	7	6	2	1	0	0	55	64
	Hardwood	0	1	2	53	4	6	1	0	0	69	77
	Scrub/Shrub	0	1	0	0	1	0	1	0	0	3	33
	Herbaceous	0	0	2	0	6	19	2	0	0	29	66
	Bare Ground	0	0	0	0	1	2	8	0	0	11	73
	Paved Roads	0	0	0	0	0	0	0	0	0	0	0
	Cantonment	0	0	0	0	0	0	0	0	0	0	0
	Column Total	0	20	42	62	19	31	13	0	0	187	
	User's Accuracy %	0.0	70	83	85	5	61	61	0	0		
Overall Accuracy = 69.5%												

POINTS OF CONTACT: For additional information, contact Mr. Scott G. Bourne, U.S. Army Engineer Research and Development Center (ERDC), Waterways Experiment Station, Vicksburg, MS (601-634-3980, Scott.Bourne@erdc.usace.army.mil). This technical note was written by Mr. Bourne and Mr. Mark R. Graves, ERDC. This document should be cited as follows:

Bourne, S. G., and Graves, M. R. (2001). "Classification of land-cover types for the Fort Benning ecoregion using enhanced thematic mapper data," *SERDP Technical Notes Collection*, ERDC/EL TN-ECMI-01-01, U.S. Army Engineer Research and Development Center, Vicksburg, MS. www.wes.army.mil/el/

ACKNOWLEDGMENTS:

This report was prepared for the Ecosystem Characterization and Monitoring Initiative (ECMI) sponsored by the Strategic Environmental Research and Development Program (SERDP) Ecosystem Management Project (SEMP). The technical monitor was Dr. Robert Holst, SERDP Program Manager.

The work was performed under the direction of the Ecosystem Evaluation and Engineering Division (EE) of the Environmental Laboratory (EL), U.S. Army Engineer Research and Development Center (ERDC). The EL Principal Investigator was Mr. Scott Bourne and co-investigator was Mr. Mark R. Graves, EL. Project Manager for the ECMI is Mr. Harold W. West, EL, and Program Manager for the SEMP is Dr. Harold E. Balbach of the Construction Engineering Research Laboratory (CERL), ERDC, Champaign, IL.

Many individuals contributed to the support of this project, including the following: Mr. John Brent, Mr. Pete Swiderek, and Ms. Theresa Davo of Fort Benning, GA, the host site for the SEMP; Drs. Rose Kress and David Price of EL, and Ms. Elizabeth Lord, Dyntel Corporation.

At the time of publication, Acting Director of EL was Dr. Edwin A. Theriot and Chief of EE was Dr. David J. Tazik. Dr. James R. Houston was Director of ERDC, and COL John W. Morris III, EN, was Commander.

REFERENCES

- Ahmad, W., Jupp, L.B., and Nunez, M. (1992). "Land cover mapping in a rugged terrain area using Landsat MSS data," *Int. J. Remote Sensing* 13(4), 673-683.
- Fleming, M. D., Berkebile, J. S., and Hoffer, R. M. (1975). "Computer aided analysis of Landsat-1 MSS data: A comparison of three approaches including a modified clustering approach," Information Note 072475, Laboratory for Application of Remote Sensing, West Lafayette, IN.
- Kress, M. R. (2000). "Executive Summary, Long-Term Monitoring Program, Fort Benning, GA, Version 2.1," Technical Report ERDC/EL TR-01-15, U.S. Army Engineer Research and Development Center, Vicksburg, MS.
- Schriever, J. R., and Congalton, R. G. (1993). "Mapping forest cover-types in New Hampshire using multi-temporal Landsat Thematic Mapper data," *ASPRS/ACSM Annual Convention & Exposition, 15-19 February (New Orleans, Louisiana)*, 3:333-342.
- Smith, C., Pyden, N., and Cole, P. (1995). "ERDAS Field Guide," 3rd ed., ERDAS, Inc., Atlanta, GA.

NOTE: *The contents of this technical note are not to be used for advertising, publication, or promotional purposes. Citation of trade names does not constitute an official endorsement or approval of the use of such products.*